

Have Sperm Densities Declined? A Reanalysis of Global Trend Data

Shanna H. Swan, Eric P. Elkin, and Laura Fenster

Reproductive Epidemiology Section, California Department of Health Services, Emeryville, CA 94608 USA

In 1992 a worldwide decline in sperm density was reported; this was quickly followed by numerous critiques and editorials. Because of the public health importance of this finding, a detailed reanalysis of data from 61 studies was warranted to resolve these issues. Multiple linear regression models (controlling for abstinence time, age, percent proven fertility, specimen collection method, study goal and location) were used to examine regional differences and the interaction between region (United States, Europe, and non-Western countries) and year. Nonlinear models and residual confounding were also examined in these data. Using a linear model (adjusted $R^2 = 0.80$), means and slopes differed significantly across regions ($p = 0.02$). Mean sperm densities were highest in Europe and lowest in non-Western countries. A decline in sperm density was seen in the United States (studies from 1938–1988; slope = -1.50 ; 95% confidence interval (CI), -1.90 – -1.10) and Europe (1971–1990; slope = -3.13 ; CI, -4.96 – -1.30), but not in non-Western countries (1978–1989; slope = 1.56 ; CI, -1.00 – 4.12). Results from nonlinear models (quadratic and spline) were similar. Thus, further analysis of these studies supports a significant decline in sperm density in the United States and Europe. Confounding and selection bias are unlikely to account for these results. However, some intraregional differences were as large as mean decline in sperm density between 1938 and 1990, and recent reports from Europe and the United States further support large interarea differences in sperm density. Identifying the cause(s) of these regional and temporal differences, whether environmental or other, is clearly warranted. **Key words:** epidemiology, multiple regression analysis, semen quality, sperm density, trend.

Environ Health Perspect 105:1228–1232 (1997). <http://ehis.niehs.nih.gov>

Among the adverse health endpoints that have been linked to endocrine-altering chemicals in the environment, male reproductive dysfunction, and particularly impaired semen quality, is of particular concern. An analysis of 61 studies of sperm density concluded that "...reports published world wide indicate clearly that sperm density has declined appreciably during 1938–1990" (1). This study stimulated considerable controversy. It was argued that fitting a linear model to these data was inappropriate (2–4) although the post-1970 increase that had been suggested was later found to be nonsignificant ($p = 0.36$) (5). Olsen et al. (2) utilized several nonlinear univariate models (spline, step function, and quadratic), which fit the data equally well and somewhat better than the linear model. These nonlinear models suggested a decline in sperm density until some time in the 1970s, at which point the slope became positive (for the spline) or the curve turned upward (for the quadratic); alternatively, the horizontal line dropped (for the step function). Thus, these models imply quite different trends in sperm density. Because of the public health importance of this question, we conducted a review of the original studies and reanalyzed these data using multiple regression methods.

Methods

Analysis of previous studies. Studies published between 1930 and 1990 that included data on sperm density were screened for eligibility

by Carlsen et al. (1). The authors' protocol excluded studies that included men in infertile couples and men referred because of genital abnormalities and studies that selected men on the basis of sperm count. Studies that used nonmanual methods for counting sperm were also excluded whenever possible, although laboratory methods were not always specified. Sixty-one studies published between 1938 and 1990 were included. [See Carlsen et al. (1) for a complete list of references.] Using linear regression, the authors found that sperm density decreased linearly during the study period at the rate of $-0.93 \times 10^6/\text{ml}/\text{year}$, decreasing from $113 \times 10^6/\text{ml}$ to $66 \times 10^6/\text{ml}$ ($p < 0.0001$).

Current analysis. We reviewed all articles cited by Carlsen et al. (1) and excluded three non-English language studies (6–8) and two studies that included men who conceived only after an infertility workup (9,10). The remaining 56 studies, including 97% of the 14,947 subjects in Carlsen et al. (1), were included in this reanalysis. Data on the following variables were abstracted and included in all multiple regression models: mean (or median) sperm density, publication year, study location, study goal, percent of men with proven fertility, semen collection method, age, and abstinence time. Variables indicating the completeness of this information were also included.

The arithmetic mean of sperm density was used when available; otherwise, the

median or geometric mean, adjusted for the difference from the arithmetic mean, was used ($n = 3$ studies). The 56 studies were stratified into regions: the United States (27 studies published 1938–1988), Europe and Australia (16 studies published 1971–1990), and other (non-Western) countries (13 studies published 1978–1989). Interaction terms to assess differences between regions in mean sperm density and slope were included in all multiple regression models. Multiple regression (using procedures for generalized linear models and regression analysis) was used to fit linear, step, spline, and quadratic models (11). It has been suggested that nonlinear models fit this data set better than the linear, because of an apparent upward turn in sperm density during the last 15–20 years of the study period (2,3). The step model assumes that mean sperm density is constant until it drops suddenly to a second constant level. The spline model fits two lines, with a possible change in slope, at a preselected time during the study period. The quadratic model includes a second-order term (year-squared), which allows for a single upward (or downward) curve in sperm density. Year was transformed for quadratic models to aid in convergence of this more complex model by using $\text{year}' = (\text{year} - 1930)$ as the independent variable. All regressions are weighted by the number of study subjects, and results are given in units of $10^6/\text{ml}$.

Because of missing information, confounding could not be completely controlled. We used information on the modeled variables, as well as additional variables abstracted from these studies, to look for evidence of residual confounding and bias. These additional variables include method of counting sperm, criterion for proven fertility, population source (sperm donors, volunteers, etc.), use of transformed sperm counts, and years of sample collection. Detailed definitions of all covariates are available upon request.

Results

Univariate linear and nonlinear models were fit to the 56 studies, and results were similar to those obtained by Olsen et al. (2) using all

Address correspondence to S.H. Swan, Reproductive Epidemiology Section, Environmental Health Investigations Branch, California Department of Health Services, 5900 Hollis Street, Suite E, Emeryville, CA 94608 USA.

The authors thank Alan Hubbard for his contribution to the statistical analysis.

Received 27 May 1997; accepted 29 July 1997.

61 studies. These are included for comparison to multiple regression models (Table 1).

Initial analysis showed significant interregional differences in mean sperm density ($p = 0.02$). A series of multiple regression models were then fit to these data. Results of the linear model (Model I) are shown in Table 2. Significant differences between regions were found for both intercepts and slopes. The greatest decline is seen for the European studies ($-3.13/\text{year}$, CI, -4.96 – -1.30). For U.S. studies, the slope is less steep ($-1.50/\text{year}$; CI, -1.90 – -1.10), but greater than that for the univariate model (-0.95). The slope for non-Western countries was positive and differed significantly from the European and U.S. slopes, although the confidence interval was wide ($1.56/\text{year}$; CI, -1.00 – 4.12). This latter group of studies did not fit any of the models well, reflecting the heterogeneity of the countries included (e.g., Brazil, India, Israel, Libya, Hong Kong, Kuwait, Nigeria, Tanzania, and Thailand), the small number of studies (12), and the short time during which these were published (12 years). Figure 1A contains the fitted regression lines for the three regions. The adjusted R^2 for the full linear model including all covariates was 0.80, compared to 0.36 for the univariate model and 0.62 for a model that only included terms for region, year, and the interaction of region and year.

Nonlinear models were also fit to these data. The quadratic model (Fig. 1D), which also fit the data well (adjusted $R^2 = 0.78$), demonstrates the absence of curvature in the U.S. studies. Some downward curvature is seen for the European studies, while means for non-Western studies show some upward curvature. However, none of the second order (year²) terms were significant in the model with regional interaction. As seen in Table 2, results for the spline model (Model II) and

linear model (Model I) were almost identical, differing only in a slight (nonsignificant) change in the U.S. slope post-1970 (from -1.52 to -1.47 ; $p = 0.97$). The similarity of these two models can be seen in Figure 1A and B. These data also fit a step function (Model III), with a significant post-1970 decrease in sperm density in all regions (see Table 2 and Figure 1C). None of these models suggest a post-1970 rise in sperm density except, possibly, in non-Western countries. The apparently improved fit of the nonlinear univariate models reported previously (2) was an artifact of confounding by region and the interaction of region and study year.

Confounding (by abstinence time, age, specimen collection method, geographic region), selection bias (changing definitions of normal men and proven fertility), measurement error (methods of counting sperm, variability in sperm counts, study year as independent variable), and statistical artifacts (choice of incorrect model, assuming sperm density normally distributed) are now explored using data from these and other studies (Table 3).

Confounding. MacLeod and Gold (13) and Magnus et al. (12) found that sperm concentration was 50–69% greater in samples collected after 10 days of abstinence than after 3 days ($p < 0.05$). Bendvold (14) reported that mean abstinence time

decreased from 7.5 to 4.4 days between 1956 and 1986. James (15) suggested that an increase in marital coital frequency may have contributed to the decline in sperm density reported in that study by shortening average abstinence time. Among the 56 studies analyzed, those with no information on abstinence times were published somewhat earlier than those with reported (or protocol-specified) abstinence time (1976 vs. 1981; $p = 0.13$). Twelve studies specified the actual mean or range of abstinence times, while 30 studies included only a protocol recommended abstinence time. However, adherence to this recommendation is uncertain. The protocol of Auger et al. (16) requested an abstinence of 3–5 days, yet only 66% of men complied. Although compliance with the recommended abstinence time was not assessed in most studies, a similar lack of compliance was likely in all studies. Therefore, because abstinence time was unknown, or known by protocol only in the majority of these studies, it could not be adequately controlled and remains a likely confounder.

The relationship between age and sperm count is complicated by the increase in abstinence time with age. In a study of 484 fertile men, Schwartz et al. (17) found that mean abstinence time increased from 3.8 days among men less than 26 years of age to

Table 1. Temporal variation in sperm density^a

Model	Adjusted R^2	Year		Spline		Step		Quadratic	
		β	p -Value	β	p -Value	β	p -Value	β	p -Value
I. Linear	0.36	-0.95	(<0.0001)						
II. Spline 1970	0.42	-1.91	(<0.0001)	2.06	(0.01)				
III. Step 1970	0.44					-36.75	(<0.0001)		
IV. Quadratic ^b	0.41	-3.32	(0.002)					0.0319	(0.02)

^aUnits for intercepts and β coefficients are $10^6/\text{ml}$.

^bFor the quadratic model, year² = year - 1930.

Table 2. Temporal and regional variation in sperm density^a

Model	Adjusted R^2	Year		Region	Region		Interaction Region \times year		Slope within region ^b
		β	p -Value		β	p -Value	β	p -Value	
I. Linear	0.80	-1.50	(<0.0001)	United States	Referent		Referent		-1.50 (-1.90–-1.10)
				Europe	3.249	(0.08)	-1.63	(0.08)	-3.13 (-4.96–-1.30)
				Non-Western	-6.070	(0.03)	3.06	(0.03)	1.56 (-1.00–4.12)
II. Spline	0.79	-1.52	(0.001)	United States before 1970	Referent		Referent		-1.52 (-2.37–-0.66)
				United States 1970 or later	—		0.04 ^c	(0.97)	-1.47 (-3.00–0.06)
				Europe	3.210	(0.05)	-1.61 ^c	(0.13)	-3.12 (-4.99–-1.26)
				Non-Western	-6.109	(0.34)	3.08 ^c	(0.04)	1.56 (-1.03–4.16)
III. Step 1970	0.72	—	—	United States before 1970	Referent		—		106.7 (91.0–122.5)
				United States 1970 or later	-39.0	(<0.0001)			67.7 (55.9–79.5)
				Europe	-31.8	(0.0003)			75.0 (60.0–90.0)
				Non-Western	-48.4	(<0.0001)			58.3 (46.0–70.7)

CI, confidence interval.

^aIn addition to publication year and region, the following covariates were included in all models: proven fertility, abstinence time, age, specimen collection method, and study goal. Units for β coefficients, slopes, and confidence intervals are $10^6/\text{ml}$.

^bValues shown are slope within region (CI) for Models I and II and mean within region (CI) for Model III.

^cChange in slope for years after 1969.

5.3 days among men over 40. Failure to adequately control for abstinence time can lead to the paradoxical finding that older men appear to have higher sperm concentrations. Nieschlag et al. (18), in a study with no control for abstinence time, found sperm concentrations over $40 \times 10^6/\text{ml}$ higher in grandfathers (mean age 67 years) than in fathers (mean age 29 years). These results may be misleading because daily sperm production has been shown to decrease with age (19). In a univariate analysis of the 42 studies that included information on age of subjects, a significantly lower sperm concentration was seen in studies that included men over 40 ($p = 0.04$). These studies were published later (mean publication year 1982) than studies in which all men were under 40 (mean year 1969) or in which age was not reported (mean year 1962); thus, age may confound this analysis.

Zavos and Goodpasture (20) found that when semen samples were obtained using a collection device during intercourse, sperm concentration was 56% higher than when samples from the same subjects were collected by masturbation ($p < 0.01$). Most of the studies we analyzed stated that semen collection was by masturbation, but some specified other methods ($n = 5$); in some studies, the collection method was unspecified ($n = 10$). Sperm concentration was significantly lower in studies with samples collected by masturbation than in the other 15 studies ($73.2 \times 10^6/\text{ml}$ vs. $99.5 \times 10^6/\text{ml}$; $p < 0.01$). Because alternative collection methods tended to be used in earlier studies (mean publication year 1956), the high concentrations in earlier studies may, at least in part, be attributable to collection methods.

Geographic region has been long recognized as an important confounder of sperm

density. In 1979, David et al. (21) compared sperm concentrations obtained in Paris (1973–1978) to those reported by Nelson and Bunge (22) from Iowa (1968–1972) and suggested that these differences ($98.4 \times 10^6/\text{ml}$ vs. $48 \times 10^6/\text{ml}$) may be due to geographic factors. Fisch and Goluboff (23) and Paulsen et al. (24) noted that trends in semen quality differed by geographic area. We compared broad geographic regions, finding significant between-region differences in slopes and intercepts. In addition, considerable variation is present within these regions. In fact, within-region differences in sperm concentration, such as those reported by Fisch et al. (25) ($131.5 \times 10^6/\text{ml}$ in New York vs. $72.7 \times 10^6/\text{ml}$ in California), are as large as the change in sperm density across the 43 years studied by Carlsen et al. (1) ($113 \times 10^6/\text{ml}$ – $66 \times 10^6/\text{ml}$).

Selection bias. Because of the nature of sample collection, no random sample of semen quality is possible and selection bias will always be of concern, though biases may differ by method of population selection. Studies of normal men may be biased by the use of an arbitrary (and changing) cutoff in sperm count to define normality (26). In 1937, Meaker (27) suggested a sperm concentration of $60 \times 10^6/\text{ml}$ as the lowest count “compatible with fertility.” This standard was lowered to $20 \times 10^6/\text{ml}$ by MacLeod and Gold (28) in 1951. Carlsen et al. (1) stratified studies, comparing “men with proven fertility” and “normal males”; we found little difference between (univariate) slopes for these groups (-0.99 vs. -1.15 ; $p = 0.74$). However, the definition of proven fertility also varied over time. Earlier studies of fertile men were more likely to require that wives be currently or recently pregnant (mean sperm density $81.4 \times 10^6/\text{ml}$; mean publication year 1972), while later studies tended to require only paternity (mean sperm density $72.7 \times$

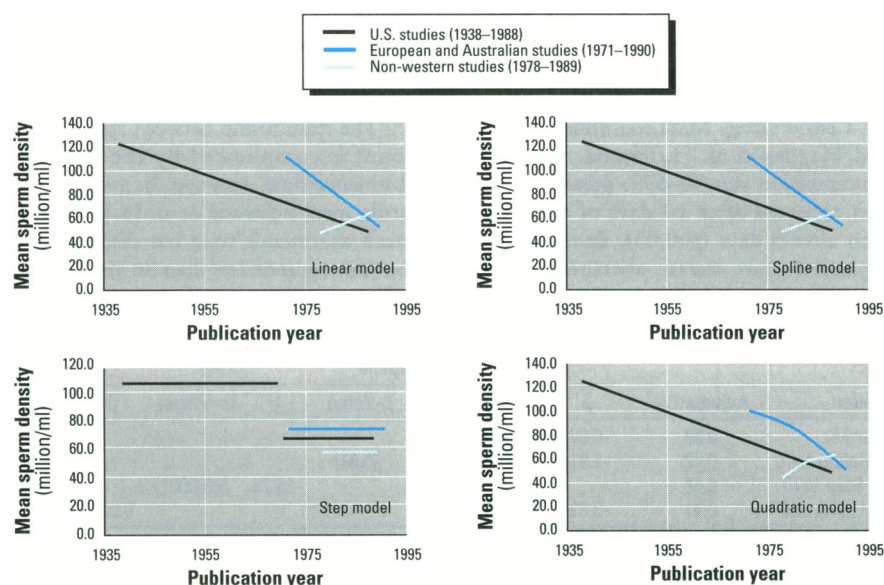


Figure 1. Four models of temporal and regional variation in sperm density controlling for proven fertility, abstinence time, age, specimen collection method, study goal, and interaction of region and study year.

Table 3. Possible sources of confounding, selection bias, measurement error, and statistical artifact in Carlsen et al. (1)

Factor	Source	How likely that factor influenced analysis of sperm density	References ^a
Confounding	Abstinence time	Likely; abstinence time increases with age, and later studies include older men	(1,2,4,23)
	Age	Possible; some evidence for increasing age over study period	(1,4,23)
	Specimen collection method	Possible; early studies allowed alternate methods	(2)
	Geographic region	Highly likely; a significant factor in modifying time trends	(23,24)
Selection bias	Changes in definition of normal men	Possible; but would not alter trend in men of proven fertility	(26)
	Changing distribution of study population	Possible; although slopes based on normal and fertile did not differ	(2,5,38)
Measurement error	Year of publication instead of collection years	Possible; interval between semen collection and publication may vary over time	(1,38)
	Variability of sperm counts	Unlikely; variability introduces random error but not bias	(2,4)
	Changes in methods of counting sperm	Unlikely; only manual methods used, and correlation between manual methods high	(2,30)
Statistical artifact	Linear model not most appropriate	Unlikely for multiple regression models; possible for univariate models	(2,3,38)
	Slope significantly increased after 1970	Unlikely; no post-1970 increase using model with interaction term	(3,5)
	Sperm counts not normally distributed	Unlikely; while true for individual counts, mean counts are approximately normal	(2,26,38)
	Mean sperm density not most appropriate	Possible; difficult to assess with few medians available for analysis	(1,2)
	Data unequally distributed over study years	Unlikely; while true, this only affects model fit and does not introduce bias	(1,2,3,5)

^aReferences to publications that discuss this factor in relation to temporal trends in sperm density.

$10^6/\text{ml}$; mean publication year 1982). Thus, trends for both groups of men may be biased by changing selection criteria.

Measurement error. Variation in sperm count, first studied in 1938 (29), is an important source of measurement error. Variation in sperm count can be partitioned into that attributable to the analytic method, to the technician, and to the subject himself. Although Carlsen et al. (1) limited their analysis to studies that used manual counting methods, even these changed during the study period (30). The laboratory protocols issued by the World Health Organization (WHO) have served to minimize this source of variability (31,32). These protocols recommend that counts be obtained by hemocytometer, as was done in most of the 56 studies analyzed, although some more recent studies used other counting devices, such as the Coulter Counter (Coulter Electronic Sales Co., Hialeah, FL). However, since the correlation between counts obtained using the hemocytometer and the Coulter Counter has been shown to be very high (0.99) (33), these changes probably had little effect on these data. Intertechnician variability is also a relatively small source of error, with a coefficient of variation estimated to be 6.1% (34). Nonetheless, authors agree that the within-subject coefficient of variation is appreciable (40–46%) (35–37). While this variability serves to decrease the precision of the trend estimates, it is not likely to introduce bias.

Use of publication year instead of sample collection years was noted by Farrow (38) as a possible source of bias. Publication year was used by Carlsen et al. (1) because collection years were seldom provided. Only nine studies included this information; all of these were published after 1980, and the interval between median collection and publication ranged from 1 to 7.5 years (mean interval 4.5 years). The effect of this error is uncertain. The selection of the independent variable for these trend analyses is a more important consideration. Whether sperm density is regressed on year of sample collection, as was done by Carlsen et al. (1), or on year of birth, as was done by Auger et al. (16) and Irvine et al. (39), depends on the hypothesis under investigation. The implications of this choice have been clarified recently by Keiding (40).

Statistical artifacts. After taking region into account using a model with interaction for region and year, the quadratic, spline, and linear models are nearly equivalent. A step function can be fit to these data, but the rationale for using it is less clear because it assumes an abrupt jump in sperm density at a single time point. All multiple regression models showed that sperm density decreased

with time; no post-1970s rise was seen, except perhaps in non-Western countries. Thus, the criticism that Carlsen et al. (1) had inappropriately fit a linear model (2–4) appears to be addressed by the use of modeling that takes region into account and separates the data from Western and non-Western countries.

Because sperm density is not normally distributed (28), the use of logarithmic or square root transformations and median (or geometric mean) rather than arithmetic mean (2,28) has been suggested. These transformations, while theoretically desirable, are not possible without access to the raw data. Moreover, such transformations are unlikely to alter these results appreciably; a univariate regression line fit to median sperm density, available in 15 studies, yielded a slope similar to that seen for mean density (slope = 1.27; $p = 0.0002$ for medians).

As discussed by Carlsen et al. (1) and others (2,3,5), studies included in this analysis were not uniformly distributed over time and space. In fact, the distribution of study publication dates and locations reflects the recent interest and technological development in this field, particularly in the United States and Europe. Eighty-five percent of subjects came from studies published between 1975 and 1990; France and the United States accounted for 71% of all subjects; and most of the world's population has not been studied at all. This unbalanced design decreases statistical power, but should not introduce bias. It may, however, limit the generalizability of the results.

Discussion

Declining semen quality and environmental causes of such a decline are not new concerns. Nelson and Bunge (22) noted sperm concentrations in 1970–1973 markedly below those reported by MacLeod and Gold (28) in 1951 and suggested that an environmental factor was responsible. Dougherty et al. (41) noted negative correlations between sperm densities and levels of organochlorines in seminal fluid. Murature et al. (42) studied sperm densities from 45 studies published 1929–1981 and found a decline in sperm concentration during 1949–1981, which was inversely correlated with several environmental exposures. James (15) concluded that sperm counts declined with publication year, at least after 1960, based on a global analysis of 29 studies of sperm concentration over a 45-year period. Conclusions from these analyses were, therefore, consistent with that of Carlsen et al. (1), despite the inclusion of different time periods and studies and the use of different exclusionary criteria.

Although the first non-U.S. study in our analysis was not published until 1971,

additional data suggest that sperm densities outside the United States were high before that time. Varnek (1944) (6) and Robles (1947) (8), included in Carlsen et al. (1) but not in this analysis (because they were published in Danish and Spanish, respectively), had mean sperm densities of $85.4 \times 10^6/\text{ml}$ and $103.2 \times 10^6/\text{ml}$, respectively. Davidson (1949) (43), not included in Carlsen although eligible, reported a mean density of $143 \times 10^6/\text{ml}$ in 15 fertile British men. The mean sperm density from five studies, published in 1944–1962, which include 2,456 infertile European men (14,44,45–47) was $98.5 \times 10^6/\text{ml}$. It is reasonable to assume that sperm densities from fertile European men would have been at least as high and therefore would have been consistent with pre-1970 data from the United States.

The multiple linear regression model shows an even steeper decline in sperm density in the United States (1938–1988) and in Europe (1971–1990) than that reported by Carlsen et al. (1), but no decline in non-Western countries, where data are sparse and are available only since 1978. When regional differences are considered, the data do not support either a “hockey stick” (spline) or an upwardly curving (quadratic) function, as has been suggested (2–4). Although confounding was controlled to the extent possible in these analyses, residual confounding may have contributed to the observed decline, but is unlikely to entirely explain it. Statistical factors (i.e., failure to transform nonnormally distributed data, use of mean vs. median, and use of publication year vs. year of sample collection) are not likely to have influenced these results appreciably. The overall downward trend does not rule out a decline in some regions and an increase (or no change) in others, nor does it rule out considerable intraregional differences, even within the United States or Europe. Recent reports from France (16,48,49), Scotland (39), Belgium (50), and the United States (23,24,51) are conflicting, supporting a decline in France, Belgium, and Scotland, but not in the United States. A recent study from London (52) suggests that differences in sperm density within a single city may be important. This between-area variation may be due to real differences between environments and populations.

This analysis demonstrates that the decline in sperm density reported by Carlsen et al. (1) is not likely to be an artifact of bias, confounding, or statistical analysis. We have not addressed the cause(s) of this decline or assumed an environmental etiology. Cross-sectional comparisons of semen quality now under way in comparably selected populations in several countries may identify areas of low (and high) sperm

concentration. Careful exposure assessment will be required to identify etiologic factors. In the future, these studies should include a broader representation of non-Western populations and banking of semen and serum to facilitate studies of biomarkers of exposure and trends in those exposures.

REFERENCES

- Carlsen E, Giwercman A, Keiding N, Skakkebaek N. Evidence for decreasing quality of semen during past 50 years. *BMJ* 305:609-613 (1992).
- Olsen GW, Bodner KM, Ramlow JM. Have sperm counts been reduced 50 percent in 50 years? A statistical model revisited. *Fertil Steril* 63:887-893 (1995).
- Brake A, Krause W. Decreasing quality of semen [letter]. *BMJ* 305:1498 (1992).
- Lerchi A, Nieschlag E. Decreasing sperm counts? A critical (re)view. *Exp Clin Endocrinol Diabetes* 104:301-307 (1996).
- Keiding N, Skakkebaek NE. Sperm decline—real or artifact? *Fertil Steril* 65:450-451 (1996).
- Varnek J. Spermaens maegde. In: *Spermaundersogesser ved sterilitet: Med specielt henblik pa spermernes morfologi*. Aarhus: Universitetsforlaget i Aarhus, 1944:42-52.
- Sturde H-C, Glowania HJ, Bohm K. Vergleichende ejakulatuntersuchungen bei man-nern aus sterilen und fertilen ehen. *Arch Derm Forsch* 241:426-437 (1971).
- Robles GG. Estudio del liquido espermatico. *Arch Peruanos Patol Clin* 1:615-661 (1947).
- Bahamondes L, Abdelmassih R, Dachs JN. Survey of 185 sperm analyses of fertile men in an infertility service. *Int J Androl* 2:526-533 (1979).
- Santomauro AG, Sciarra JJ, Varma AO. A clinical investigation of the role of the semen analysis and postcoital test in the evaluation of male infertility. *Fertil Steril* 23:245-251 (1972).
- SAS Institute Inc. *SAS/STAT User's Guide* 6, Vol 2. 4th Ed. Cary, NC: SAS Institute Inc, 1989.
- Magnus O, Tollefsrud A, Abyholm T, Purvis K. Effects of varying the abstinence period in the same individuals on sperm quality. *Arch Androl* 26:199-203 (1991).
- MacLeod J, Gold RZ. The male factor in fertility and infertility. V. Effect of continence on semen quality. *Fertil Steril* 3:297-315 (1952).
- Bendvold E. Semen quality in Norwegian men over a 20-year period. *Int J Fertil* 34:401-404 (1989).
- James WH. Secular trend in reported sperm counts. *Andrologia* 12:381-388 (1980).
- Auger J, Kunstmann JM, Czyglik F, Jouannet P. Decline in semen quality among fertile men in Paris during the past 20 years. *N Engl J Med* 332:281-285 (1995).
- Schwartz D, Mayaux MJ, Spira A, Moscato L, Jouannet P, Czyglik F, David G. Study of a group of 484 fertile men: Part II: Relation between age (20-59) and semen characteristics. *Int J Androl* 4:450-456 (1981).
- Nieschlag E, Lammers U, Freischem CW, Langer K, Wickings EJ. Reproductive functions in young fathers and grandfathers. *J Clin Endocrinol Metab* 55:676-681 (1982).
- Johnson L. Spermatogenesis and aging in the human. *J Androl* 7:331-354 (1986).
- Zavos PM, Goodpasture JC. Clinical improvements of specific seminal deficiencies via intercourse with a seminal collection device versus masturbation. *Fertil Steril* 51:190-193 (1989).
- David G, Jouannet P, Martin-Boyce A, Spira A, Schwartz D. Sperm counts in fertile and infertile men. *Fertil Steril* 31:453-455 (1979).
- Nelson MK, Bunge RG. Semen analysis: evidence for changing parameters of male fertility potential. *Fertil Steril* 25:503-507 (1974).
- Fisch H, Goluboff ET. Geographic variations in sperm counts: a potential cause of bias in studies of semen quality. *Fertil Steril* 65:1044-1046 (1996).
- Paulsen CA, Berman NG, Wang C. Data from men in greater Seattle area reveals no downward trend in semen quality: further evidence that deterioration of semen quality is not geographically uniform. *Fertil Steril* 65:1015-1020 (1996).
- Fisch H, Goluboff ET, Olson JH, Feldshuh J, Broder SJ, Barad DH. Semen analysis in 1,283 men from the United States over a 25-year period: no decline in quality. *Fertil Steril* 65:1009-1014 (1996).
- Bromwich P, Cohen J, Stewart I, Walker A. Decline in sperm counts: an artifact of changed reference range of "normal"? *BMJ* 309:19-22 (1994).
- Meaker SR. *Human Sterility*. Baltimore, MD: Williams & Wilkins, 1943.
- MacLeod J, Gold RZ. The male factor in fertility and infertility. II. Spermatozoon counts in 1000 men of known fertility and in 1000 cases of infertile marriage. *J Urol* 66:436-449 (1951).
- Hotchkiss RS. Factors in stability and variability of semen specimens. *J Urol* 45:875-888 (1941).
- Forti G, Serio M. Male infertility; is its rising incidence due to better methods of detection or an increasing frequency? *Hum Reprod* 8:1153-1154 (1993).
- World Health Organization. *Manual for the Examination of Human Semen and Semen-Human-Cervical Mucus Interaction* (Belsey MA, Eliasson R, Gallegos AJ, Moghissi KS, Paulsen CA, Prasad MRN, eds). Singapore: Press Concern, 1980.
- World Health Organization. *WHO Laboratory Manual for the Examination of Human Semen and Semen-Cervical Mucus Interaction*. Cambridge, UK: Cambridge University Press, 1987.
- Brotherton J, Barnard G. Estimation of number, mean size and size distribution of human spermatozoa in oligospermia using a Coulter counter. *J Reprod Fertil* 40:341-357 (1974).
- Cooper TG, Neuwinger J, Bahrs S, Nieschlag E. Internal quality control of semen analysis. *Fertil Steril* 58:172-178 (1992).
- MacLeod J, Gold RZ. Male factor in fertility and infertility. VIII. A study of variation in semen quality. *Fertil Steril* 7:387-410 (1956).
- Schrader SM, Turner TW, Breitenstein MJ, Simon SD. Longitudinal study of semen quality of unexposed workers. I. Study overview. *Reprod Toxicol* 2:183-190 (1988).
- Mallidis C, Howard EJ, Baker HWG. Variation of semen quality in normal men. *Int J Androl* 14:99-107 (1991).
- Farrow S. Falling sperm quality: fact or fiction? *BMJ* 309:1-2 (1994).
- Irvine S, Cawood E, Richardson D, MacDonald E, Aitken J. Evidence of deteriorating semen quality in the United Kingdom: birth cohort study in 577 men in Scotland over 11 years. *BMJ* 312:467-470 (1996).
- Keiding N. Assessing effects of age, period and cohort on human sperm concentration. In: *Liber Amicorum Roel van Strik*. Rotterdam: University of Rotterdam, 1996:45-48.
- Dougherty RC, Whitaker MJ, Tang S-Y, Bottcher R, Keller M, Kuehl DW. In: *Environmental Health Chemistry: The Chemistry of Environmental Agents as Potential Human Hazards* (McKinney JD, ed). Ann Arbor, MI: Ann Arbor Science Publication 1981:263-278.
- Murature DA, Tang SY, Steinhart G, Dougherty RC. Phthalate esters and semen quality parameters. *Biomed Environ Mass Spectrom* 14:473-477 (1987).
- Davidson HA. Male subfertility: interim report of 3,182 cases. *BMJ* 2:1328-1332 (1949).
- Hammen R. Studies on impaired fertility in man: with special reference to the male. Copenhagen: Einar Munksgaard, 1944.
- Bostofte E, Serup J, Rebbe H. Has the fertility of Danish men declined through the years in terms of semen quality? A comparison of semen qualities between 1952 and 1972. *Int J Fertil* 28:91-95 (1983).
- Osser S, Liedholm P, Ranstam J. Depressed semen quality: a study over two decades. *Arch Androl* 12:113-116 (1984).
- Bendvold E, Gortlieb C, Bygdeman M, Eneroth P. Depressed semen quality in Swedish men from barren couples: a study over three decades. *Arch Androl* 26:189-194 (1991).
- deMouzon J, Thonneau P, Spira A, Multigner L. Declining sperm count. Semen quality has declined among men born in France since 1950 [letter]. *BMJ* 313:43 (1996).
- Bujan L, Mansat A, Pontonnier F, Mieusset R. Time series analysis of sperm concentration in fertile men in Toulouse, France between 1977 and 1992. *BMJ* 312:471-472 (1996).
- Van Waelegheem K, De Clercq N, Vermeulen L, Schoonjans F, Comhaire F. Deterioration of sperm quality in young healthy Belgian men. *Hum Reprod* 11:325-329 (1996).
- Wittmaak FM, Shapiro SS. Longitudinal study of semen quality in Wisconsin men over one decade. *Wisconsin Med J* 91:477-479 (1992).
- Ginsburg J, Okolo S, Prelevic G, Hardiman P. Residence in the London area and sperm density [letter]. *Lancet* 343:230 (1994).